

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Horvitz, Robert
Yuan, Junying
Shaham, Shai

(ii) TITLE OF THE INVENTION: RELATEDNESS OF HUMAN INTERLEUKIN-1
BETA CONVERTASE GENE TO A C. ELEGANS CELL DEATH
GENE, INHIBITORY PORTIONS OF THESE GENES AND...

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Clark & Elbing LLP
(B) STREET: 176 Federal Street
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02110

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE: 21-MAY-1995
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 08/394,189
(B) FILING DATE: 24-FEB-1995

(A) APPLICATION NUMBER: 08/282,211
(B) FILING DATE: 12-JUL-1994

(A) APPLICATION NUMBER: 07/984,182
(B) FILING DATE: 20-NOV-1992

(A) APPLICATION NUMBER: 07/897,788
(B) FILING DATE: 12-JUN-1992

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Bieker-Brady, Kristina
(B) REGISTRATION NUMBER: 39,109

	TAAAAATAAC	GATTTCTCAT	TGAAAATTGT	GTTTTATGTT	TGCGAAAATA	AAAGAGAACT	1920
	GATTCAAAAC	AATTTTAACA	AAAAAAAACC	CCAAAATTCG	CCAGAAATCA	AGATAAAAAA	1980
	TTCAAGAGGG	TCAAAATTTT	CCGATTTTAC	TGACTTTCAC	CTTTTTTTTC	GTAGTTCAGT	2040
	GCAGTTGTTG	GAGTTTTTGA	CGAAAAGTAG	GAAAAAATC	GATAAAAATT	ACTCAAATCG	2100
5	AGCTGAATTT	TGAGGACAAT	GTTTAAAAAA	AAACACTATT	TTTCCAATAA	TTTCACTCAT	2160
	TTTCAGACTA	AATCGAAAAT	CAAATCGTAC	TCTGACTACG	GGTCAGTAGA	GAGGTCAACC	2220
	ATCAGCCGAA	GATGATGCGT	CAAGATAGAA	GGAGCTTGCT	AGAGAGGAAC	ATTATGATGT	2280
	TCTCTAGTCA	TCTAAAAGTC	GATGAAATTY	TCGAAGTTCT	CATCGCAAAA	CAAGTGTTGA	2340
	ATAGTGATAA	TGGAGATATG	ATTAATGTGA	GTTTTTAATC	GAATAATAAT	TTTAAAAAAA	2400
10	AATTGATAAT	ATAAAGAATA	TTTTTGCAGT	CATGTGGAAC	GGTTCGCGAG	AAGAGACGGG	2460
	AGATCGTGAA	AGCAGTGCAA	CGACGGRGAG	ATGTGGCGTT	CGACGCGTTT	TATGATGCTC	2520
	TTTCGCTCTAC	GGGACACGAA	GGACTTGCTG	AAGTTCCTGA	ACCTCTCGCC	AGATCGTAGG	2580
	TTTTTAAAGT	TCGGCGCAAA	AGCAAGGGTC	TCACGGAAAA	AAGAGGCGGA	TCGTAATTTT	2640
	GCAACCCACC	GGCACGGTTT	TTTCTCCGA	AAATCGGAAA	TTATGCACTT	TCCCAAATAT	2700
15	TTGAAGTGAA	ATATATTTTA	TTTACTGAAA	GCTCGAGTGA	TTATTTATTT	TTTAAACACTA	2760
	ATTTTCGTGG	CGCAAAAGGC	CATTTTGTAG	ATTTGCCGAA	AATACTTGTC	ACACACACAC	2820
	ACACACATCT	CCTTCAAATA	TCCCTTTTTT	CAGTGTTGAC	TCGAATGCTG	TCGAATTCGA	2880
	GTGTCCAATG	TCACCGGCAA	GCCATCGTCG	GAGCCGCGCA	TTGAGCCCCG	CCGGCTACAC	2940
	TTCACCGACC	CGAGTTCACC	GTGACAGCGT	CTCTTCAGTG	TCATCATTCA	CTTCTTATCA	3000
20	GGATATCTAC	TCAAGAGCAA	GATCTCGTTC	TCGATCGCGT	GCACTTCATT	CATCGGATCG	3060
	ACACAATTAT	TCATCTCCTC	CAGTCAACGC	ATTTCCCAGC	CAACCTTGTA	TGTTGATGCG	3120
	AACACTAAAT	TCTGAGAATG	CGCATTACTC	AACATATTTG	ACGCGCAAAT	ATCTCGTAGC	3180
	GAAAAATACA	GTAACCCCTT	AAATGACTAT	TGTAGTGTG	ATTTACGGGC	TCGATTTTCG	3240
	AAACGAATAT	ATGCTCGAAT	TGTGACAACG	AATTTTAATT	TGTCATTTTT	GTGTTTTCTT	3300
25	TTGATATTTT	TGATCAATTA	ATAAATTATT	TCCGTAAACA	GACACCAGCG	CTACAGTACT	3360
	CTTTTAAAGA	GTTACAGTAG	TTTTTCGCTT	AAGATATTTT	GAAAAGAATT	TTAAACATTT	3420
	TGAAAAAATA	TCATCTAACA	TGTGCCAAAA	CGCTTTTTTT	AAGTTTCGCA	GATTTTTTGA	3480
	TTTTTTTCAT	TCAAGATATG	CTTATTAACA	CATATAATTA	TCATTAATGT	GAATTTCTTG	3540
	TAGAAATTTT	GGGCTTTTTG	TTCTAGTATG	CTCTACTTTT	GAAATTGCTC	AACGAAAAAA	3600
30	TCATGTGGTT	TGTTTCATATG	AATGACGAAA	AATAGCAATT	TTTTATATAT	TTTCCCCTAT	3660
	TCATGTTGTG	CAGAAAAATA	GTAAAAAAGC	GCATGCATTT	TTTCGACATTT	TTTACATCGA	3720
	ACGACAGCTC	ACTTCACATG	CTGAAGACGA	GAGACGCGGA	GAAATACCAC	ACATCTTTCT	3780
	GCGTCTCTCG	TCTTCAGCAT	GTGAAATGGG	ATCTCGGTCTG	ATGTAAAAAA	ATGTCGAATA	3840
	ATGTAAAAAA	TGCATGCGTT	TTTTTACACT	TTTCTGCACA	AATGAATAGG	GGGAAAATGT	3900
35	ATTAATAATAC	ATTTTTTGTA	TTTTTCAACA	TCACATGATT	AACCCCATTA	TTTTTTCGTT	3960
	GAGCAACTTA	AAAAGTAGAG	AATATTAGAG	CGAAAACCAA	AATTTCTTCA	AGATATTACC	4020
	TTTATTGATA	ATTATAGATG	TTAATAAGCA	TATCTTGAAT	GAAAGTCAGC	AAAAATATGT	4080
	GCGAAACACC	TGAAAAAAT	CAAAAATTCT	GCGAAAATTG	AAAAAATGCA	TTAAATATACA	4140
	TTTTTGCATT	TTTCTACATC	ACATGAATGT	AGAAAATTAA	AAGGGAAATC	AAAATTTCTA	4200
40	GAGGATATAA	TTGAATGAAA	CATTGCGAAA	TTAAATGTG	CGAAACGTCA	AAAAAGAGGA	4260
	AATTTGGGTA	TCAAAATCGA	TCCTAAAACC	AACACATTTT	AGCATCCGCC	AACCTTTCAT	4320
	TCACCGGATG	CTCTTCTCTC	GGATACAGTT	CAAGTCGTAA	TCGCTCATT	AGCAAAGCTT	4380
	CTGGACCAAC	TCAATACATA	TTCCATGAAG	AGGATATGAA	CTTTGTGCGT	GCACCAACCA	4440
	TAAGCCGTGT	TTTCGACGAG	AAAACCATGT	ACAGAAACTT	CTCGAGTCCT	CGTGGAATGT	4500
45	GCCTCATCAT	AAATAATGAA	CACTTTGAGC	AGATGCCAAC	ACGGAATGGT	ACCAAGGCCG	4560
	ACAAGGACAA	TCTTACCAAT	TTGTTTCAGAT	GCATGGGCTA	TACGGTTATT	TGCAAGGACA	4620
	ATCTGACGGG	AAGGGTACGG	CGAAATTATA	TTACCCAAAC	GCGAAATTTG	CCATTTTGCG	4680
	CCGAAAATGT	GGCGCCCGGT	CTCGACACGA	CAATTTGTGT	TAAATGCAAA	AATGTATAAT	4740
	TTTGCAAAAA	ACAAAATTTT	GAACCTCCGC	GAAAATGATT	TACCTAGTTT	CGAAATTTTC	4800
50	GTTTTTTCCG	GCTACATTAT	GTGTTTTTTC	TTAGTTTTTTC	TATAATATTT	GATGTAAAAA	4860
	ACCGTTTGTA	AATTTTCAGA	CAATTTTCCG	CATACAAAAC	TTGATAGCAC	GAAATCAATT	4920

	TTCTGAATTT	TCAAAATTAT	CCAAAAATGC	ACAATTTAAA	ATTTGTGAAA	ATTGGCAAAC	4980
	GGTGTTTCAA	TATGAAATGT	ATTTTTTAAA	ACTTTAAAAA	CCACTCCGGA	AAAGCAATAA	5040
	AAATCAAAAAC	AACGTCACAA	TTCAAATTCA	AAAGTTATTC	ATCCGATTTG	TTTATTTTTG	5100
	CAAAATTTGA	AAAAATCATG	AAGGATTTAG	AAAAGTTTTA	TAACATTTTT	TCTAGATTTT	5160
5	TCAAAATTTT	TTTAAACAAA	TCGAGAAAAA	GAGAATGAAA	AATCGATTTT	AAAAATATCC	5220
	ACAGCTTTCGA	GAGTTTGAAA	TTACAGTACT	CCTTAAAGGC	GCACACCCCA	TTTGCATTGG	5280
	ACCAAAAAAT	TGTCGTGTCT	AGACCAGGTA	CCGTAGTTTT	TGTCGCAAAA	ATTGCACCAT	5340
	TGGACAATAA	ACCTTCCTAA	TCACCAAAAA	GTAAAAATTGA	AATCTTCGAA	AAGCCAAAAA	5400
	ATTCAAAAAA	AAAGTCGAAT	TTCGATTTTT	TTTTTGTTTT	TTTGGTCCCA	AAAACCAAAA	5460
10	AAATCAATTT	TCTGCAAAAT	ACCAAAAAGA	AACCCGAAAA	AATTTCCCAG	CCTTGTTCCCT	5520
	AATGTAAACT	GATATTTAAT	TTCCAGGGAA	TGCTCCTGAC	AATTCGAGAC	TTTGCCAAAC	5580
	ACGAATCACA	CGGAGATTCT	GCGATACTCG	TGATTCTATC	ACACGGAGAA	GAGAATGTGA	5640
	TTATTGGAGT	TGATGATATA	CCGATTAGTA	CACACGAGAT	ATATGATCTT	CTCAACGCGG	5700
	CAAATGCTCC	CCGTCTGGCG	AATAAGCCGA	AAATCGTTTT	TGTGCAGGCT	TGTCGARGCG	5760
15	GTTCTGTTTT	TATTTTAATT	TTAATATAAA	TATTTTAAAT	AAATTCATTT	TCAGAACGTC	5820
	GTGACAATGG	ATTCCCAGTC	TTGGATTCTG	TCGACGGAGT	TCCTGCATTT	CCTTCGTCGTG	5880
	GATGGGACAA	TCGAGACGGG	CCATTGTTCA	ATTTTCTTGG	ATGTGTGCGG	CCGCAAGTTY	5940
	AGGTTGCAAT	TTAATTTCTT	GAATGAGAAT	ATTCTTCAA	AAAAATCTAA	ATAGATTTTT	6000
	ATTCCAGAAA	GTCCCGATCG	AAAAATTGCG	ATATAATTAC	GAAATTTGTG	ATAAAATGAC	6060
20	AAACCAATCA	GCATCGTCGA	TCTCCGCCCA	CTTCATCGGA	TTGGTTTGAA	AGTGGGCGGA	6120
	GTGAATTGCT	GATTGGTCGC	AGTTTTCAGT	TTAGAGGGAA	TTAAAAATC	GCCTTTTCGA	6180
	AAATTAAAAA	TTGATTTTTT	CAATTTTTTC	GAAAAATATT	CCGATTATTT	TATATTCTTT	6240
	GGAGCGAAAG	CCCCGTCTCT	TAAACATTTT	TAAATGATAA	TTAATAAATT	TTTGCARCAA	6300
	GTGTGGAGAA	AGAAGCCGAG	CYAAGCTGAC	ATTCTGATTC	GRTACGCAAC	GACAGCTCAA	6360
25	TATGTTTTCGT	GGAGAAACAG	TGCTCGTGGA	TCATGGTTCA	TTCAAGCCGT	CTGTGAAGTG	6420
	TTCTCGACAC	ACGYAAAGGA	TATGGATGTT	GTTGAGCTGC	TGACTGAAGT	CAATAAGAAG	6480
	GTCGYTTGTG	GATTTTCAGAC	ATCACAGGGA	TCGAATATTT	TGAAACAGAT	GCCARAGGTA	6540
	CTTGAAACAA	ACAATGCATG	TCTAACTTTT	AAGGACACAG	AAAAATAGGC	AGAGGCTCCT	6600
	TTTGCAAGCC	TGCCGCGCGT	CAACCTAGAA	TTTTAGTTTT	TAGCTAAAAT	GATTGATTTT	6660
30	GAATATTTTA	TGCTAATTTT	TTTGCGTTAA	ATTTTGAAAT	AGTCACTATT	TATCGGGTTT	6720
	CCAGTAAAAA	ATGTTTATTA	GCCATTGGAT	TTTACTGAAA	ACGAAAATTT	GTAGTTTTTC	6780
	AACGAAATTT	ATCGATTTTT	AAATGTAAAA	AAAAATAGCG	AAAAATTACAT	CAACCATCAA	6840
	GCATTTAAGC	CAAAATTGTT	AACTCATTTA	AAAATTAAAT	CAAAGTTGTC	CACGAGTATT	6900
	ACACGGTTGG	CGCGCGGCAA	GTTTGCAAAA	CGACGCTCCG	CCTCTTTTTT	TGTGCGGCTT	6960
35	GAAAACAAGG	GATCGGTTTA	GATTTTTCCC	CAAAATTTAA	ATTAAATTTT	AGATGACATY	7020
	CCGCCTGCTC	AAAAAGTTCT	ACTTTTGGCC	GGAAGCACGA	AACTCTGCCG	TCTAAAATTC	7080
	ACTCGTGATT	CATTGCCCAA	TTGATAATTG	TCTGTATCTT	CTCCCCAGT	TCTCTTTCGC	7140
	CCAATTAGTT	TAAAACCATG	TGTATATTGT	TATCCTATAC	TCATTTCACT	TTATCATTCT	7200
	ATCATTTCTC	TTCCCATTTT	CACACATTTT	CATTTCTCTA	CGATAATCTA	AAATTATGAC	7260
40	GTTTGTGTCT	CGAACGCATA	ATAATTTTAA	TAACTCGTTT	TGAATTTGAT	TAGTTGTTGT	7320
	GCCCAGTATA	TATGTATGTA	CTATGCTTCT	ATCAACAAAA	TAGTTTCATA	GATCATCACC	7380
	CCAACCCAC	CAACCTACCG	TACCATATTC	ATTTTGGCCG	GGAATCAATT	TCGATTAATT	7440
	TTAACCTATT	TTTTCGCCAC	AAAAAATCTA	ATATTTGAAT	TAACGAATAG	CATTCCCATC	7500
	TCTCCCCTGC	CGGAATGCCT	CCCGGCCTTT	TAAAGTTCGG	AACATTTGGC	AATTATGTAT	7560
45	AAATTTGTAG	GTCCCCCCCC	TCAATTTCCG	CCCATCATCT	CAAATTGCAT	TCTTTTTTCG	7620
	CCGTGATATC	CCGATTCTGG	TCAGCAAAGA	TCT			7653

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 503 amino acids

(B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

5

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

	Met	Met	Arg	Gln	Asp	Arg	Arg	Ser	Leu	Leu	Glu	Arg	Asn	Ile	Met	Met
	1				5					10					15	
10	Phe	Ser	Ser	His	Leu	Lys	Val	Asp	Glu	Ile	Xaa	Glu	Val	Leu	Ile	Ala
				20					25					30		
	Lys	Gln	Val	Leu	Asn	Ser	Asp	Asn	Gly	Asp	Met	Ile	Asn	Ser	Cys	Gly
			35					40					45			
	Thr	Val	Arg	Glu	Lys	Arg	Arg	Glu	Ile	Val	Lys	Ala	Val	Gln	Arg	Arg
		50					55					60				
15	Xaa	Asp	Val	Ala	Phe	Asp	Ala	Phe	Tyr	Asp	Ala	Leu	Arg	Ser	Thr	Gly
	65					70					75				80	
	His	Glu	Gly	Leu	Ala	Glu	Val	Leu	Glu	Pro	Leu	Ala	Arg	Ser	Val	Asp
				85						90					95	
	Ser	Asn	Ala	Val	Glu	Phe	Glu	Cys	Pro	Met	Ser	Pro	Ala	Ser	His	Arg
20				100						105					110	
	Arg	Ser	Arg	Ala	Leu	Ser	Pro	Ala	Gly	Tyr	Thr	Ser	Pro	Thr	Arg	Val
			115					120					125			
	His	Arg	Asp	Ser	Val	Ser	Ser	Val	Ser	Ser	Phe	Thr	Ser	Tyr	Gln	Asp
		130				135						140				
25	Ile	Tyr	Ser	Arg	Ala	Arg	Ser	Arg	Ser	Arg	Ser	Arg	Ala	Leu	His	Ser
	145					150					155				160	
	Ser	Asp	Arg	His	Asn	Tyr	Ser	Ser	Pro	Pro	Val	Asn	Ala	Phe	Pro	Ser
				165						170					175	
	Gln	Pro	Ser	Ser	Ala	Asn	Ser	Ser	Phe	Thr	Gly	Cys	Ser	Ser	Leu	Gly
30				180					185						190	
	Tyr	Ser	Ser	Ser	Arg	Asn	Arg	Ser	Phe	Ser	Lys	Ala	Ser	Gly	Pro	Thr
			195					200					205			
	Gln	Tyr	Ile	Phe	His	Glu	Glu	Asp	Met	Asn	Phe	Val	Asp	Ala	Pro	Thr
		210					215					220				
35	Ile	Ser	Arg	Val	Phe	Asp	Glu	Lys	Thr	Met	Tyr	Arg	Asn	Phe	Ser	Ser
	225					230					235				240	
	Pro	Arg	Gly	Met	Cys	Leu	Ile	Ile	Asn	Asn	Glu	His	Phe	Glu	Gln	Met
				245						250					255	
	Pro	Thr	Arg	Asn	Gly	Thr	Lys	Ala	Asp	Lys	Asp	Asn	Leu	Thr	Asn	Leu
40				260					265					270		
	Phe	Arg	Cys	Met	Gly	Tyr	Thr	Val	Ile	Cys	Lys	Asp	Asn	Leu	Thr	Gly
			275					280				285				
	Arg	Gly	Met	Leu	Leu	Thr	Ile	Arg	Asp	Phe	Ala	Lys	His	Glu	Ser	His
		290				295					300					
45	Gly	Asp	Ser	Ala	Ile	Leu	Val	Ile	Leu	Ser	His	Gly	Glu	Glu	Asn	Val
	305					310					315				320	
	Ile	Ile	Gly	Val	Asp	Asp	Ile	Pro	Ile	Ser	Thr	His	Glu	Ile	Tyr	Asp
				325					330						335	

10988824 062201

```

Leu Leu Asn Ala Ala Asn Ala Pro Arg Leu Ala Asn Lys Pro Lys Ile
      340                      345                      350
Val Phe Val Gln Ala Cys Arg Xaa Glu Arg Arg Asp Asn Gly Phe Pro
      355                      360                      365
5  Val Leu Asp Ser Val Asp Gly Val Pro Ala Phe Leu Arg Arg Gly Trp
      370                      375                      380
Asp Asn Arg Asp Gly Pro Leu Phe Asn Phe Leu Gly Cys Val Arg Pro
385                      390                      395                      400
10 Gln Val Xaa Gln Val Trp Arg Lys Lys Pro Ser Xaa Ala Asp Ile Leu
      405                      410                      415
Ile Arg Tyr Ala Thr Thr Ala Gln Tyr Val Ser Xaa Arg Asn Ser Ala
      420                      425                      430
Arg Gly Ser Trp Phe Ile Gln Ala Val Cys Glu Val Phe Ser Thr His
      435                      440                      445
15 Xaa Lys Asp Met Asp Val Val Glu Leu Leu Thr Glu Val Asn Lys Lys
      450                      455                      460
Val Xaa Cys Gly Phe Gln Thr Ser Gln Gly Ser Asn Ile Leu Lys Gln
465                      470                      475                      480
20 Met Pro Xaa Met Thr Xaa Arg Leu Leu Lys Lys Phe Tyr Phe Trp Pro
      485                      490                      495
Glu Ala Arg Asn Ser Ala Val
      500

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1373 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 18...1229

(D) OTHER INFORMATION: /product= "human interleukin-1 beta convertase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

AAAAGGAGAG AAAAGCC ATG GCC GAC AAG GTC CTG AAG GAG AAG AGA AAG      50
      Met Ala Asp Lys Val Leu Lys Glu Lys Arg Lys
      1                      5                      10
40 CTG TTT ATC CGT TCC ATG GGT GAA GGT ACA ATA AAT GGC TTA CTG GAT      98
Leu Phe Ile Arg Ser Met Gly Glu Gly Thr Ile Asn Gly Leu Leu Asp
      15                      20                      25
GAA TTA TTA CAG ACA AGG GTG CTG AAC AAG GAA GAG ATG GAG AAA GTA      146

```

	Glu	Leu	Leu	Gln	Thr	Arg	Val	Leu	Asn	Lys	Glu	Glu	Met	Glu	Lys	Val		
	30						35				40							
5	AAA Lys	CGT Arg	GAA Glu	AAT Asn	GCT Ala	ACA Thr	GTT Val	ATG Met	GAT Asp	AAG Lys	ACC Thr	CGA Arg	GCT Ala	TTG Leu	ATT Ile	GAC Asp	194	
	45						50				55							
	TCC Ser	GTT Val	ATT Ile	CCG Pro	AAA Lys	GGG Gly	GCA Ala	CAG Gln	GCA Ala	TGC Cys	CAA Gln	ATT Ile	TGC Cys	ATC Ile	ACA Thr	TAC Tyr	242	
	60					65				70						75		
10	ATT Ile	TGT Cys	GAA Glu	GAA Glu	GAC Asp	AGT Ser	TAC Tyr	CTG Leu	GCA Ala	GGG Gly	ACG Thr	CTG Leu	GGA Gly	CTC Leu	TCA Ser	GCA Ala	290	
					80				85				90					
	GAT Asp	CAA Gln	ACA Thr	TCT Ser	GGA Gly	AAT Asn	TAC Tyr	CTT Leu	AAT Asn	ATG Met	CAA Gln	GAC Asp	TCT Ser	CAA Gln	GGA Gly	GTA Val	338	
				95				100				105						
15	CTT Leu	TCT Ser	TCC Ser	TTT Phe	CCA Pro	GCT Ala	CCT Pro	CAG Gln	GCA Ala	GTG Val	CAG Gln	GAC Asp	AAC Asn	CCA Pro	GCT Ala	ATG Met	386	
	110						115				120							
20	CCC Pro	ACA Thr	TCC Ser	TCA Ser	GGC Gly	TCA Ser	GAA Glu	GGG Gly	AAT Asn	GTC Val	AAG Lys	CTT Leu	TGC Cys	TCC Ser	CTA Leu	GAA Glu	434	
	125						130				135							
	GAA Glu	GCT Ala	CAA Gln	AGG Arg	ATA Ile	TGG Trp	AAA Lys	CAA Gln	AAG Lys	TCG Ser	GCA Ala	GAG Glu	ATT Ile	TAT Tyr	CCA Pro	ATA Ile	482	
	140					145				150						155		
25	ATG Met	GAC Asp	AAG Lys	TCA Ser	AGC Ser	CGC Arg	ACA Thr	CGT Arg	CTT Leu	GCT Ala	CTC Leu	ATT Ile	ATC Ile	TGC Cys	AAT Asn	GAA Glu	530	
				160				165				170						
	GAA Glu	TTT Phe	GAC Asp	AGT Ser	ATT Ile	CCT Pro	AGA Arg	AGA Arg	ACT Thr	GGA Gly	GCT Ala	GAG Glu	GTT Val	GAC Asp	ATC Ile	ACA Thr	578	
				175				180				185						
30	GGC Gly	ATG Met	ACA Thr	ATG Met	CTG Leu	CTA Leu	CAA Gln	AAT Asn	CTG Leu	GGG Gly	TAC Tyr	AGC Ser	GTA Val	GAT Asp	GTG Val	AAA Lys	626	
	190						195				200							
35	AAA Lys	AAT Asn	CTC Leu	ACT Thr	GCT Ala	TCG Ser	GAC Asp	ATG Met	ACT Thr	ACA Thr	GAG Glu	CTG Leu	GAG Glu	GCA Ala	TTT Phe	GCA Ala	674	
	205						210				215							
	CAC His	CGC Arg	CCA Pro	GAG Glu	CAC His	AAG Lys	ACC Thr	TCT Ser	GAC Asp	AGC Ser	ACG Thr	TTC Phe	CTG Leu	GTG Val	TTC Phe	ATG Met	722	
	220					225				230						235		

[illegible]

(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ala Asp Lys Val Leu Lys Glu Lys Arg Lys Leu Phe Ile Arg Ser
1 5 10 15
Met Gly Glu Gly Thr Ile Asn Gly Leu Xaa Asp Glu Leu Leu Gln Thr
10 20 25 30
Arg Val Leu Asn Lys Glu Glu Met Glu Lys Val Lys Arg Glu Asn Ala
35 40 45
Thr Val Met Asp Lys Thr Arg Ala Leu Ile Asp Ser Val Ile Pro Lys
50 55 60
15 Xaa Ala Gln Ala Cys Gln Ile Cys Ile Thr Tyr Ile Cys Glu Glu Asp
65 70 75 80
Ser Tyr Leu Ala Gly Thr Leu Gly Leu Ser Ala Asp Gln Thr Ser Gly
85 90 95
Asn Tyr Leu Asn Met Gln Asp Ser Gln Gly Val Leu Ser Ser Phe Pro
100 105 110
20 Ala Pro Gln Ala Val Gln Asp Asn Pro Ala Met Pro Thr Ser Ser Gly
115 120 125
Ser Glu Gly Asn Val Lys Leu Cys Ser Leu Glu Glu Ala Gln Arg Ile
130 135 140
25 Trp Lys Gln Lys Ser Ala Glu Ile Tyr Pro Ile Met Asp Lys Ser Ser
145 150 155 160
Arg Thr Arg Leu Ala Leu Ile Ile Cys Asn Glu Glu Phe Asp Ser Ile
165 170 175
Pro Arg Arg Thr Gly Ala Glu Val Asp Ile Thr Gly Met Thr Met Leu
180 185 190
30 Leu Gln Asn Leu Gly Tyr Ser Val Asp Val Lys Lys Asn Leu Thr Ala
195 200 205
Ser Asp Met Thr Thr Glu Leu Glu Ala Phe Ala His Arg Pro Glu His
210 215 220
35 Lys Thr Ser Asp Ser Thr Phe Leu Val Phe Met Ser His Gly Ile Arg
225 230 235 240
Glu Gly Ile Cys Gly Lys Lys His Ser Glu Gln Val Pro Asp Ile Leu
245 250 255
Gln Leu Asn Ala Ile Phe Asn Met Leu Asn Thr Lys Asn Cys Pro Ser
260 265 270
40 Leu Lys Asp Lys Pro Lys Val Ile Ile Ile Gln Ala Xaa Arg Xaa Asp
275 280 285
Ser Pro Gly Val Val Trp Phe Lys Asp Ser Val Gly Val Ser Gly Asn
290 295 300
45 Leu Ser Leu Pro Thr Thr Glu Glu Phe Glu Asp Asp Ala Ile Lys Lys
305 310 315 320
Ala His Ile Xaa Lys Asp Phe Ile Ala Phe Cys Ser Ser Thr Pro Asp
325 330 335

FO2290" E428850

Asn Val Ser Xaa Arg His Pro Thr Met Gly Ser Val Phe Ile Gly Arg
340 345 350
Leu Ile Glu His Met Gln Glu Tyr Xaa Cys Ser Cys Asp Val Glu Glu
355 360 365
5 Ile Phe Arg Lys Val Arg Phe Ser Phe Glu Gln Pro Asp Gly Arg Ala
370 375 380
Gln Met Pro Thr Thr Xaa Arg Val Xaa Leu Thr Arg Cys Phe Tyr Leu
385 390 395 400
Phe Pro Gly His

10

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 505 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

15

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

20

Met Met Arg Gln Asp Arg Trp Leu Leu Glu Arg Asn Ile Leu Glu Phe
1 5 10 15
Ser Ser Lys Leu Gln Ala Asp Leu Ile Leu Asp Val Leu Ile Ala Lys
20 25 30
Gln Val Leu Asn Ser Asp Asn Gly Asp Val Ile Asn Ser Cys Arg Thr
35 40 45
25 Glu Arg Asp Asn Glu Lys Glu Ile Val Lys Ala Val Gln Arg Arg Gly
50 55 60
Asp Glu Ala Phe Asp Ala Phe Tyr Asp Ala Leu Arg Asp Thr Gly His
65 70 75 80
Asn Asp Leu Ala Asp Val Leu Met Pro Leu Ser Arg Pro Xaa Xaa Xaa
85 90 95
30 Asn Pro Val Pro Met Glu Cys Pro Met Ser Pro Ser Ser His Arg Arg
100 105 110
Ser Arg Ala Leu Ser Pro Pro Xaa Tyr Ala Ser Pro Thr Arg Val His
115 120 125
35 Arg Asp Ser Ile Ser Ser Val Ser Ser Phe Thr Ser Thr Tyr Gln Asp
130 135 140
Val Tyr Ser Arg Ala Arg Ser Ser Ser Arg Ser Ser Arg Pro Leu Gln
145 150 155 160
Ser Ser Asp Arg His Asn Tyr Met Ser Ala Ala Thr Ser Phe Pro Ser
165 170 175
40 Gln Pro Xaa Ser Ala Asn Ser Ser Phe Thr Gly Cys Ala Ser Leu Gly
180 185 190
Tyr Ser Ser Ser Arg Asn Arg Ser Phe Ser Lys Thr Ser Ala Gln Ser
195 200 205
45 Gln Tyr Ile Phe His Glu Glu Asp Met Asn Tyr Val Asp Ala Pro Thr
210 215 220

	Asn	Gly	Asp	Met	Ile	Asn	Ser	Cys	Arg	Thr	Glu	Arg	Asp	Asn	Glu	Lys	
			20						25					30			
	Glu	Ile	Val	Lys	Ala	Val	Gln	Arg	Gly	Asp	Glu	Ala	Phe	Asp	Ala		
			35					40					45				
5	Phe	Tyr	Asp	Ala	Leu	Arg	Asp	Thr	Gly	His	Asn	Asp	Leu	Ala	Asp	Val	
		50					55					60					
	Leu	Met	Pro	Leu	Ser	Arg	Pro	Val	Asp	Ser	Asn	Pro	Val	Pro	Met	Glu	
	65					70					75					80	
	Cys	Pro	Met	Ser	Pro	Ser	Ser	His	Arg	Arg	Ser	Arg	Ala	Leu	Ser	Pro	
					85					90					95		
10	Pro	Xaa	Tyr	Ala	Ser	Pro	Thr	Arg	Val	His	Arg	Asp	Ser	Ile	Ser	Ser	
				100					105					110			
	Val	Ser	Ser	Phe	Thr	Ser	Thr	Tyr	Gln	Asp	Val	Tyr	Ser	Arg	Ala	Thr	
			115					120					125				
15	Ser	Ser	Ser	Pro	Leu	Gln	Thr	Ser	Asp	Arg	His	Asn	Tyr	Val	Ser	Ala	
		130					135					140					
	Ser	Thr	Ser	Phe	Gln	Ser	Gln	Pro	Ala	Ser	Ala	Asn	Ser	Ser	Phe	Thr	
	145					150					155					160	
	Gly	Ser	Ala	Ser	Leu	Gly	Tyr	Ser	Ser	Ser	Arg	Thr	Arg	Ser	Tyr	Ser	
					165					170					175		
20	Lys	Thr	Ser	Ala	His	Ser	Gln	Tyr	Ile	Phe	His	Glu	Glu	Asp	Met	Asn	
			180							185				190			
	Tyr	Val	Asp	Ala	Pro	Thr	Ile	His	Arg	Val	Phe	Asp	Glu	Lys	Thr	Met	
		195						200					205				
25	Tyr	Arg	Asn	Phe	Ser	Thr	Pro	Arg	Gly	Leu	Cys	Leu	Ile	Ile	Asn	Asn	
		210					215					220					
	Glu	His	Phe	Glu	Gln	Met	Pro	Thr	Arg	Asn	Gly	Thr	Lys	Pro	Asp	Lys	
	225					230					235					240	
	Asp	Asn	Ile	Ser	Asn	Ile	Phe	Arg	Cys	Met	Gly	Tyr	Ile	Val	His	Cys	
					245					250					255		
30	Lys	Asp	Asn	Leu	Thr	Gly	Arg	Glu	Met	Met	Ser	Thr	Ile	Arg	Ser	Phe	
			260						265					270			
	Gly	Arg	Asn	Asp	Thr	His	Gly	Asp	Ser	Ala	Ile	Leu	Val	Ile	Leu	Ser	
			275					280					285				
35	His	Gly	Glu	Xaa	Asn	Val	Ile	Ile	Gly	Val	Asp	Asp	Val	Ser	Val	Asn	
		290					295					300					
	Val	His	Glu	Ile	Tyr	Xaa	Leu	Leu	Asn	Ala	Ala	Asn	Ala	Pro	Arg	Leu	
	305					310					315					320	
	Ala	Asn	Lys	Pro	Lys	Leu	Val	Phe	Val	Gln	Ala	Cys	Arg	Gly	Glu	Arg	
					325					330					335		

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

5 GGTTTAAACC AGTTACTCAA G

21

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

10

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCGGTGACAT TGGACACTC

19

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

ACTATTCAAC ACTTG

15

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 171 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

25

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

(D) OTHER INFORMATION: Xaa at position 117 is Ala or Val.

30

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Leu Thr Val Gln Val Tyr Arg Thr Ser Gln Lys Cys Ser Ser Ser
1 5 10 15

Lys His Val Val Glu Val Leu Leu Asp Pro Leu Gly Thr Ser Phe Cys
 20 25 30
 Ser Leu Leu Pro Pro Pro Leu Leu Leu Tyr Glu Thr Asp Arg Gly Val
 35 40 45
 5 Asp Gln Gln Asp Gly Lys Asn His Thr Gln Ser Pro Gly Cys Glu Glu
 50 55 60
 Ser Asp Ala Gly Lys Glu Glu Leu Met Lys Met Arg Leu Pro Thr Arg
 65 70 75 80
 Ser Asp Met Ile Cys Gly Tyr Ala Cys Leu Lys Gly Asn Ala Ala Met
 85 90 95
 10 Arg Asn Thr Lys Arg Gly Ser Trp Tyr Ile Glu Ala Leu Thr Gln Val
 100 105 110
 Phe Ser Glu Arg Xaa Cys Asp Met His Val Ala Asp Met Leu Val Lys
 115 120 125
 15 Val Asn Ala Leu Ile Lys Glu Arg Glu Gly Tyr Ala Pro Gly Thr Glu
 130 135 140
 Phe His Arg Cys Lys Glu Met Ser Glu Tyr Cys Ser Thr Leu Cys Gln
 145 150 155 160
 Gln Leu Tyr Leu Phe Pro Gly Tyr Pro Pro Thr
 165 170
 20

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 402 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Met Ala Asp Lys Ile Leu Arg Ala Lys Arg Lys Gln Phe Ile Asn Ser
 1 5 10 15
 Val Ser Ile Gly Thr Ile Asn Gly Leu Leu Asp Glu Leu Leu Glu Lys
 20 25 30
 Arg Val Leu Asn Gln Glu Glu Met Asp Lys Ile Lys Leu Ala Asn Ile
 35 35 40 45
 Thr Ala Met Asp Lys Ala Arg Asp Leu Cys Asp His Val Ser Lys Lys
 50 55 60
 Gly Pro Gln Ala Ser Gln Ile Phe Ile Thr Tyr Ile Cys Asn Glu Asp
 65 70 75 80
 Cys Tyr Leu Ala Gly Ile Leu Glu Leu Gln Ser Ala Pro Ser Ala Glu
 85 90 95
 40 Thr Phe Val Ala Thr Glu Asp Ser Lys Gly Gly His Pro Ser Ser Ser
 100 105 110
 Glu Thr Lys Glu Glu Gln Asn Lys Glu Asp Gly Thr Phe Pro Gly Leu
 115 120 125
 45 Thr Gly Thr Leu Lys Phe Cys Pro Leu Glu Lys Ala Gln Lys Leu Trp
 130 135 140

Lys Glu Asn Pro Ser Glu Ile Tyr Pro Ile Met Asn Thr Thr Thr Arg
 145 150 155 160
 Thr Arg Leu Ala Leu Ile Ile Cys Asn Thr Glu Phe Gln His Leu Ser
 165 170 175
 5 Pro Arg Val Gly Ala Gln Val Asp Leu Arg Glu Met Lys Leu Leu Leu
 180 185 190
 Glu Asp Leu Gly Tyr Thr Val Lys Val Lys Glu Asn Leu Thr Ala Leu
 195 200 205
 10 Glu Met Val Lys Glu Val Lys Glu Phe Ala Ala Cys Pro Glu His Lys
 210 215 220
 Thr Ser Asp Ser Thr Phe Leu Val Phe Met Ser His Gly Ile Gln Glu
 225 230 235 240
 Gly Ile Cys Gly Thr Thr Tyr Ser Asn Glu Val Ser Asp Ile Leu Lys
 245 250 255
 15 Val Asp Thr Ile Phe Gln Met Met Asn Thr Leu Lys Cys Pro Ser Leu
 260 265 270
 Lys Asp Lys Pro Lys Val Ile Ile Ile Gln Ala Cys Arg Gly Glu Lys
 275 280 285
 20 Gln Gly Val Val Leu Leu Lys Asp Ser Val Arg Asp Ser Glu Glu Asp
 290 295 300
 Phe Leu Thr Asp Ala Ile Phe Glu Asp Asp Gly Ile Lys Lys Ala His
 305 310 315 320
 Ile Glu Lys Asp Phe Ile Ala Phe Cys Ser Ser Thr Pro Asp Asn Val
 325 330 335
 25 Ser Trp Arg His Pro Val Arg Gly Ser Leu Phe Ile Glu Ser Leu Ile
 340 345 350
 Lys His Met Lys Glu Tyr Ala Trp Ser Cys Asp Leu Glu Asp Ile Phe
 355 360 365
 30 Arg Lys Val Arg Phe Ser Phe Glu Gln Pro Glu Phe Arg Leu Gln Met
 370 375 380
 Pro Thr Ala Asp Arg Val Thr Leu Thr Lys Arg Phe Tyr Leu Phe Pro
 385 390 395 400
 Gly His

35

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

40

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

(D) OTHER INFORMATION: Xaa at position 3 is Ala, His, Gln, Lys, Phe, Cha or Asp.

45

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Tyr Val Xaa Asp

1

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- 10 (D) OTHER INFORMATION: Peptide has an acetyl group at the amino terminus and an aldehyde group at the carboxyl terminus.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Tyr Val Ala Asp

1

15 (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Tyr Val Ala Asp

1

25 (2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- 30 (A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- 35 (D) OTHER INFORMATION: Xaa at position 2 is D-Ala. Peptide has an acetyl group at the amino terminus and an aldehyde group at the carboxyl terminus.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Tyr Xaa Ala Asp

1

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

5

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

10

(ix) FEATURE:

(D) OTHER INFORMATION: Peptide has an acetyl group at the amino terminus and an aldehyde group at the carboxyl terminus.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Tyr Val Lys Asp

15

1

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

20

- (A) LENGTH: 354 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: C-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

25

Arg Ser Arg Ser Arg Ser Arg Ala Leu His Ser Ser Asp Arg His Asn
1 5 10 15

Tyr Ser Ser Pro Pro Val Asn Ala Phe Pro Ser Gln Pro Ser Ser Ala
20 25 30

30

Asn Ser Ser Phe Thr Gly Cys Ser Ser Leu Gly Tyr Ser Ser Ser Arg
35 40 45

Asn Arg Ser Phe Ser Lys Ala Ser Gly Pro Thr Gln Tyr Ile Phe His
50 55 60

Glu Glu Asp Met Asn Phe Val Asp Ala Pro Thr Ile Ser Arg Val Phe
65 70 75 80

35

Asp Glu Lys Thr Met Tyr Arg Asn Phe Ser Ser Pro Arg Gly Met Cys
85 90 95

Leu Ile Ile Asn Asn Glu His Phe Glu Gln Met Pro Thr Arg Asn Gly
100 105 110

40

Thr Lys Ala Asp Lys Asp Asn Leu Thr Asn Leu Phe Arg Cys Met Gly
115 120 125

Tyr Thr Val Ile Cys Lys Asp Asn Leu Thr Gly Arg Gly Met Leu Leu

[illegible]

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 131 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Val	Asp	Gly	Val	Pro	Ala	Phe	Leu	Arg	Arg	Gly	Trp	Asp	Asn	Arg	Asp
1			5					10						15	
Gly	Pro	Leu	Phe	Asn	Phe	Leu	Gly	Cys	Val	Arg	Pro	Gln	Val	Xaa	Gln
			20					25					30		
Val	Trp	Arg	Lys	Lys	Pro	Ser	Xaa	Ala	Asp	Ile	Leu	Ile	Arg	Tyr	Ala
			35				40					45			
Thr	Thr	Ala	Gln	Tyr	Val	Ser	Xaa	Arg	Asn	Ser	Ala	Arg	Gly	Ser	Trp
	50					55					60				

Phe Ile Gln Ala Val Cys Glu Val Phe Ser Thr His Xaa Lys Asp Met
 65 70 75 80
 Asp Val Val Glu Leu Thr Glu Val Asn Lys Lys Val Xaa Cys Gly
 85 90 95
 5 Phe Gln Thr Ser Gln Gly Ser Asn Ile Leu Lys Gln Met Pro Xaa Met
 100 105 110
 Thr Xaa Arg Leu Leu Lys Lys Phe Tyr Phe Trp Pro Glu Ala Arg Asn
 115 120 125
 Ser Ala Val
 10 130

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 223 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

20 Arg Ser Arg Ser Arg Ser Arg Ala Leu His Ser Ser Asp Arg His Asn
 1 5 10 15
 Tyr Ser Ser Pro Pro Val Asn Ala Phe Pro Ser Gln Pro Ser Ser Ala
 20 25 30
 Asn Ser Ser Phe Thr Gly Cys Ser Ser Leu Gly Tyr Ser Ser Ser Arg
 25 35 40 45
 Asn Arg Ser Phe Ser Lys Ala Ser Gly Pro Thr Gln Tyr Ile Phe His
 50 55 60
 Glu Glu Asp Met Asn Phe Val Asp Ala Pro Thr Ile Ser Arg Val Phe
 65 70 75 80
 30 Asp Glu Lys Thr Met Tyr Arg Asn Phe Ser Ser Pro Arg Gly Met Cys
 85 90 95
 Leu Ile Ile Asn Asn Glu His Phe Glu Gln Met Pro Thr Arg Asn Gly
 100 105 110
 Thr Lys Ala Asp Lys Asp Asn Leu Thr Asn Leu Phe Arg Cys Met Gly
 115 120 125
 35 Tyr Thr Val Ile Cys Lys Asp Asn Leu Thr Gly Arg Gly Met Leu Leu
 130 135 140
 Thr Ile Arg Asp Phe Ala Lys His Glu Ser His Gly Asp Ser Ala Ile
 145 150 155 160
 40 Leu Val Ile Leu Ser His Gly Glu Glu Asn Val Ile Ile Gly Val Asp
 165 170 175
 Asp Ile Pro Ile Ser Thr His Glu Ile Tyr Asp Leu Leu Asn Ala Ala
 180 185 190
 Asn Ala Pro Arg Leu Ala Asn Lys Pro Lys Ile Val Phe Val Gln Ala
 195 200 205
 45 Cys Arg Xaa Glu Arg Arg Asp Asn Gly Phe Pro Val Leu Asp Ser

210

215

220

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 294 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Phe Pro Ala Pro Gln Ala Val Gln Asp Asn Pro Ala Met Pro Thr Ser
 1 5 10 15
 Ser Gly Ser Glu Gly Asn Val Lys Leu Cys Ser Leu Glu Glu Ala Gln
 20 25 30
 Arg Ile Trp Lys Gln Lys Ser Ala Glu Ile Tyr Pro Ile Met Asp Lys
 35 40 45
 Ser Ser Arg Thr Arg Leu Ala Leu Ile Ile Cys Asn Glu Glu Phe Asp
 50 55 60
 Ser Ile Pro Arg Arg Thr Gly Ala Glu Val Asp Ile Thr Gly Met Thr
 65 70 75 80
 Met Leu Leu Gln Asn Leu Gly Tyr Ser Val Asp Val Lys Lys Asn Leu
 85 90 95
 Thr Ala Ser Asp Met Thr Thr Glu Leu Glu Ala Phe Ala His Arg Pro
 100 105 110
 Glu His Lys Thr Ser Asp Ser Thr Phe Leu Val Phe Met Ser His Gly
 115 120 125
 Ile Arg Glu Gly Ile Cys Gly Lys Lys His Ser Glu Gln Val Pro Asp
 130 135 140
 Ile Leu Gln Leu Asn Ala Ile Phe Asn Met Leu Asn Thr Lys Asn Cys
 145 150 155 160
 Pro Ser Leu Lys Asp Lys Pro Lys Val Ile Ile Ile Gln Ala Xaa Arg
 165 170 175
 Xaa Asp Ser Pro Gly Val Val Trp Phe Lys Asp Ser Val Gly Val Ser
 180 185 190
 Gly Asn Leu Ser Leu Pro Thr Thr Glu Glu Phe Glu Asp Asp Ala Ile
 195 200 205
 Lys Lys Ala His Ile Xaa Lys Asp Phe Ile Ala Phe Cys Ser Ser Thr
 210 215 220
 Pro Asp Asn Val Ser Xaa Arg His Pro Thr Met Gly Ser Val Phe Ile
 225 230 235 240
 Gly Arg Leu Ile Glu His Met Gln Glu Tyr Xaa Cys Ser Cys Asp Val
 245 250 255
 Glu Glu Ile Phe Arg Lys Val Arg Phe Ser Phe Glu Gln Pro Asp Gly
 260 265 270
 Arg Ala Gln Met Pro Thr Thr Xaa Arg Val Xaa Leu Thr Arg Cys Phe
 275 280 285

Tyr Leu Phe Pro Gly His
290

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 107 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Ser Val Gly Val Ser Gly Asn Leu Ser Leu Pro Thr Thr Glu Glu Phe
1 5 10 15
Glu Asp Asp Ala Ile Lys Lys Ala His Ile Xaa Lys Asp Phe Ile Ala
20 25 30
Phe Cys Ser Ser Thr Pro Asp Asn Val Ser Xaa Arg His Pro Thr Met
35 40 45
Gly Ser Val Phe Ile Gly Arg Leu Ile Glu His Met Gln Glu Tyr Xaa
50 55 60
Cys Ser Cys Asp Val Glu Glu Ile Phe Arg Lys Val Arg Phe Ser Phe
65 70 75 80
Glu Gln Pro Asp Gly Arg Ala Gln Met Pro Thr Thr Xaa Arg Val Xaa
85 90 95
Leu Thr Arg Cys Phe Tyr Leu Phe Pro Gly His
100 105

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 187 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Phe Pro Ala Pro Gln Ala Val Gln Asp Asn Pro Ala Met Pro Thr Ser
1 5 10 15
Ser Gly Ser Glu Gly Asn Val Lys Leu Cys Ser Leu Glu Glu Ala Gln
20 25 30
Arg Ile Trp Lys Gln Lys Ser Ala Glu Ile Tyr Pro Ile Met Asp Lys
35 40 45
Ser Ser Arg Thr Arg Leu Ala Leu Ile Ile Cys Asn Glu Glu Phe Asp

	50		55		60														
	Ser	Ile	Pro	Arg	Arg	Thr	Gly	Ala	Glu	Val	Asp	Ile	Thr	Gly	Met	Thr			
	65					70					75					80			
5	Met	Leu	Leu	Gln	Asn	Leu	Gly	Tyr	Ser	Val	Asp	Val	Lys	Lys	Asn	Leu			
				85						90					95				
	Thr	Ala	Ser	Asp	Met	Thr	Thr	Glu	Leu	Glu	Ala	Phe	Ala	His	Arg	Pro			
				100					105					110					
	Glu	His	Lys	Thr	Ser	Asp	Ser	Thr	Phe	Leu	Val	Phe	Met	Ser	His	Gly			
			115					120					125						
10	Ile	Arg	Glu	Gly	Ile	Cys	Gly	Lys	Lys	His	Ser	Glu	Gln	Val	Pro	Asp			
		130					135					140							
	Ile	Leu	Gln	Leu	Asn	Ala	Ile	Phe	Asn	Met	Leu	Asn	Thr	Lys	Asn	Cys			
	145				150					155						160			
	Pro	Ser	Leu	Lys	Asp	Lys	Pro	Lys	Val	Ile	Ile	Ile	Gln	Ala	Xaa	Arg			
15					165					170					175				
	Xaa	Asp	Ser	Pro	Gly	Val	Val	Trp	Phe	Lys	Asp								
				180					185										

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 172 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

- 25 (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

	Met	Leu	Thr	Val	Gln	Val	Tyr	Arg	Thr	Ser	Gln	Lys	Cys	Ser	Ser	Ser			
	1				5					10					15				
	Lys	His	Val	Val	Glu	Val	Leu	Leu	Asp	Pro	Leu	Gly	Thr	Ser	Phe	Cys			
30			20					25					30						
	Ser	Leu	Leu	Pro	Pro	Pro	Leu	Leu	Leu	Tyr	Glu	Thr	Asp	Arg	Gly	Val			
			35				40					45							
	Asp	Gln	Gln	Asp	Gly	Lys	Asn	His	Thr	Gln	Ser	Pro	Gly	Cys	Glu	Glu			
		50					55					60							
35	Ser	Asp	Ala	Gly	Lys	Glu	Glu	Leu	Met	Lys	Met	Arg	Leu	Pro	Thr	Arg			
	65				70					75						80			
	Ser	Asp	Met	Ile	Cys	Gly	Tyr	Ala	Cys	Leu	Lys	Gly	Asn	Ala	Ala	Met			
				85						90					95				
	Arg	Asn	Thr	Lys	Arg	Gly	Ser	Trp	Tyr	Ile	Glu	Ala	Leu	Thr	Gln	Val			
40			100						105					110					
	Phe	Ser	Glu	Arg	Ala	Xaa	Cys	Asp	Met	His	Val	Ala	Asp	Met	Leu	Val			
			115					120					125						
	Lys	Val	Asn	Ala	Leu	Ile	Lys	Glu	Arg	Glu	Gly	Tyr	Ala	Pro	Gly	Thr			
		130					135					140							
45	Glu	Phe	His	Arg	Cys	Lys	Glu	Met	Ser	Glu	Tyr	Cys	Ser	Thr	Leu	Cys			

